A Monitoring Program for the Desert Tortoise

David R. Anderson and Kenneth P. Burnham
Colorado Cooperative Fish & Wildlife Research Unit
Room 201 Wagar Building, Colorado State University
Fort Collins, CO 80525-1484

Comprehensive survey methods are needed to monitor population trends in desert tortoise (*Gopherus agassizii*) in the 6 Recovery Units and allow comparisons with areas outside these Units. Such surveys will require substantial money and personnel and a strong managerial commitment to continue the monitoring for many years. A key issue in the design of survey methods is the variability, in time and space, in the proportion of tortoises that are above ground during the survey periods (e.g., in some years there will be few individuals above ground for long periods of time, even in the spring, when surveys are to be conducted). A second consideration is the fact the tortoise are difficult to detect, somewhat spatially aggregated, and are relatively rare, thus making surveying tedious. A third consideration is the very large size of both reserved (i.e., Recovery Units) and non-reserved (surrounding) lands.

Three primary objectives are: (1) intensive monitoring to detect possible, short-term, drastic declines in population density, (2) less intensive, long-term monitoring to detect possible long-term increases in population density due to management alternatives on the reserved lands, and (3) compare trends across reserved and surrounding lands. The management alternatives being applied to reserved lands are designed to lead to population increases and the eventual delisting of the species.

This document is a slight revision of one distributed at a 2-day conference on this issue held in Laughlin, NV, November 20-21, 1996. The main change here is the addition of a section on the estimated power to detect various trends; these results lead us to consider the initial, intensive monitoring for 5 years (instead of 3-4 suggested earlier). Other changes were for clarification of various levels of stratification, the change in individual transect lengths from 4.8 to 4 km, and some further thoughts on estimation of the population sex ratio.

1. Overview

We suggest a survey/monitoring program using two independent teams of observers each year, one using line transect sampling (Buckland et al. 1993) and the other using radio telemetry:

Survey Team A: A stratified system of line transects, using many, relatively short, lines. Beginning points for each line transect within a stratum would be drawn either randomly or systematically with a random start and the initial location of each transect found on the ground using a GPS. Some areas within each stratum could be deleted for both sampling and inference concerning average density (e.g., cities, airfields, permanent lakes, high mountains).

The same lines would be used each year (i.e., permanent lines, permanently marked and marked well). Each "line" would actually be a square (see figure, below) this would minimize walking time while not on a transect. Each leg of the square would be 1,000m, for a total of 4 km for that line (l_i). Such lines would serve as the basic sampling unit. Lines would be surveyed by a 2 person crew, covering 1 line transect per day in the field.

Approximately a dozen crews will be required to obtain the required sample size and complete the survey in a relatively short period of time (about 40 days). Surveys would be done in parts of April and May each year. All tortoises detected would be recorded (see example field form, attached), however, the estimation of population density would likely relate to tortoises > 160 or 180 mm in length. Burrows must be searched for tortoises (see below). Location of each tortoise detected would be made using a GPS and recorded. Each tortoise detected would be given a permanent, numbered mark or tag and released at the point where it was detected.
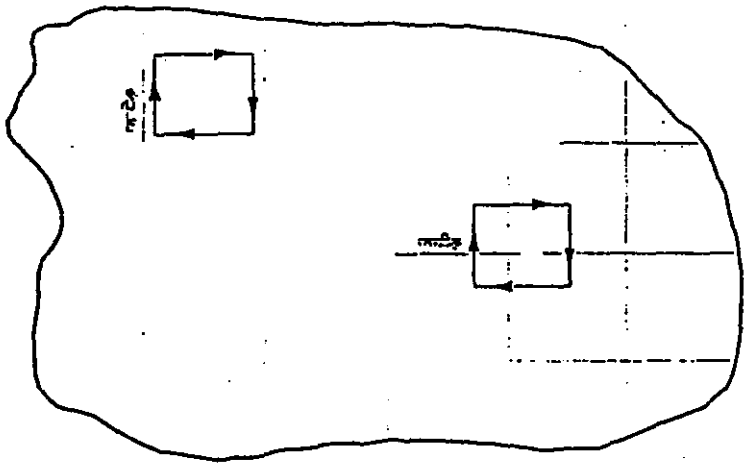
Survey Team B: A stratified survey using radio-marked tortoises to estimate the proportion of tortoises above ground. Call this proportion g_0; of course, it would vary by strata, and year (as well as within the actual survey period). At the strata level, within a year, g_0 is only crudely approximately as a binomial random variable with sampling variance $\left(g_0(1-g_0)\right)/n_r$, where n_r = number of tortoise with radio attached); improved estimates of the variability would employ some empirical estimation. Sex and size of tortoise would be recorded as well as their location (above or below ground) during this survey. Specific field protocol for the estimation of g_0 remains to be developed.

Stratification would be the same on both surveys and Team B would be sampling in the same areas as Team A. Efforts of the two survey teams would have to be closely coordinated in both sampling time and geographically within a year. Several levels of stratification are suggested:

The primary geographic strata would be the 6 Recovery Units – these are the basis for potential delisting.

The secondary geographic strata would be the 14 Desert Wildlife Management Areas (DWMAs).

Within each of the secondary geographic strata there would be some further sub-stratification (perhaps 2-4), bringing the total number of sub-strata to 30-50 and perhaps, eventually, as many as 70. Such sub-strata provide efficiency in the estimation for the 2 levels of geographic strata; density estimates for these sub-strata are of relatively little interest in themselves. Stratification would allow estimates of density or abundance for individual stratum levels as well as the total area.

Surveys A and B would be concurrent in time and area and closely integrated. The relative effort (and, thus, costs) of the two types of survey teams would be dictated by the relative variance components. Pooling over the 6 primary strata, we might hope to achieve a cv of 10-15% for the entire surveyed area; precision on the individual, primary strata would be substantially less (cv of perhaps 20-40%).

In the following material we will try to outline an overall survey/monitoring program. We provide some details that some people might want to skip over on the first reading. The initial focus is on the survey/monitoring program for the reserved areas; sampling for areas outside these areas and comparative analysis methods remain to be addressed (in general, we would recommend the same type of stratified sampling, using line transects and radio telemetry). Several other subjects need considerably more thought (e.g., exact methodologies for Survey Team B, total line length required and, thus, sample size). This is a working document that allows others to refine the suggestions made. Readers must be somewhat familiar with distance sampling (Buckland et al. 1993); however, we will provide some introductory material to help those not exposed to these methods. The first several introductory sections (below) relate to a single survey area, without the complications of stratification.

2. Critical Assumptions

The unbiased estimation of tortoise density using line transect sampling rests on two critical assumptions:

1. All tortoises on the centerline are detected with certainty. This applies to those that are on the surface (the ones in burrows are separately estimated – see material below). Tortoises some distance away from the transect centerline may be missed (for a wide variety of reasons), however, those on the line are assumed to be detected. Thus, field protocol must takes steps to assure the validity of this assumption. For example, observers should attempt to walk transect legs traveling uphill, walking more slowly when cover is dense, and looking diligently into vegetation or rock piles on or near the centerline.

2. Perpendicular distances (x) from the line to each detected tortoise are measured accurately along a line of length l. If the centerline of each transect is clearly marked and perpendicular distances are measured with a steel tape, this assumptions should be easy to meet in the field.

[A third assumption deals with movement (in relationship to the approaching observer) of animals prior to detection; this seems relatively unimportant in surveying tortoises.]

3. Estimation of the Detection Function and P_a

Central to the concept of distance sampling is the <u>detection function</u> $g(x)$:

$g(x)$ = the probability of detecting an object, given that it is at distance x from the random line

= prob{detection | distance x}.

The x_i are measured only to w, the transect <u>width</u>. Within a sub-stratum, the individual lines l_i are summed for a total transect length of L.

The shape of the detection function "evolves" as the survey data are collected and reflects the 4 factors that contribute to the n tortoises detected during the survey:

(1) tortoise density,
(2) environmental variables (e.g., habitat and soil types, effects of rainfall),
(3) observer variables (e.g., interest, training, fatigue), and
(4) variables related to the individual tortoise (e.g., size, coloration, position).

Buckland et al. (1993) provide a robust theory for the estimation of the detection function and program DISTANCE implements these methods. In summary, given the above two assumptions, the detection function can be estimated from the n perpendicular distances $(x_1, x_2, ..., x_n)$ that are measured accurately during the survey.

In line transect sampling, only a proportion of the tortoises in the area surveyed (of size $2wL$) is detected. One of the major advantages of line transect sampling is that it allows tortoises to be missed on sampling units and still provide unbiased estimates of density. In many terrestrial surveys, 60-80% of the objects of interest are not detected in the surveyed strip of size $2wL$; still unbiased estimates of density can be made, given the two assumptions. Let this unknown proportion of tortoises that are detected be denoted as P_a. In fact, given the two assumptions (see above) P_a can be estimated from the distance data using the relationship,

$$P_a = \frac{\int_0^w g(x)\, dx}{w}\ .$$

Program DISTANCE (Laake et al. 1994) allows the estimation of the detection function from the distance data (i.e., $\hat{g}(x)$), computes the integral, and provides an objective estimate of P_a and its standard error.

In strip transect sampling, one must blindly assume that $P_a \equiv 1$; this cannot be verified unless detection distances are measured. Strip transect sampling is a special case of line transect sampling (i.e., when all objects within the strip are detected with certainty). In fact, program DISTANCE allows the estimation of density from strip transects by forcing $g(x) \equiv 1$ for all distances between the line and the strip width (w).

4. Estimation of the Proportion of Tortoises that are on the Surface During the Survey, g_0

The proportion of tortoises above ground (denoted as g_0) is highly variable across time (i.e., days and weeks during survey periods, in addition to annual changes) and space, and much of this variation is correlated with precipitation, temperature, and other environmental variables. During a survey we suspect that some burrows are not detected and that not all tortoises in burrows are found, even if their burrow is detected. This is an important issue, varies geographically, and must received adequate attention. Substantial data on the proportion of tortoises above ground already exists and this information would be helpful in survey design. For the first few years, annual estimates of g_0 would be made for each level of stratification, from monitoring telemetered tortoises.

An operational definition of "above ground" might be those tortoise seen on the surface or those in burrows that can be seen by eye with only the aid of mirrors or lights held outside the mouth of the burrow. Thus, tortoise deeper in burrows would not be

recorded by Survey Team A. A clear definition here is essential so that the two survey teams are not operating under two overlapping sets of rules.

The long-term goal here would be to "understand" why tortoises remain on the surface vs. going into burrows (during the survey periods). Thus, a goal would be to *model* g_0 as a function of several covariates, say Y_i (obviously, temperature and precipitation, size, sex, etc.). A lot is known about how to proceed here – transform the simple proportion g_0 to

$$\text{logit}(g_0) = \log_e\left(g_0/(1-g_0)\right)$$

and then use a linear model of the environmental covariates) as,

$$\text{logit}(g_0) = \beta_0 + \beta_1(Y_1) + \beta_2(Y_2) + \cdots + \beta_m(Y_m),$$

where the Y_j represent the m covariates. A good model of g_0 will depend on substantial variation in the covariates and this will take several years. Estimates of g_0 must be made by Survey Team B during the entire time the populations are being surveyed by Survey Team A using line transect sampling. Estimation of g_0 is obtained by back-transforming,

$$\hat{g}_0 = \left(1 + \exp(-(\hat{\beta}_0 + \hat{\beta}_1(Y_1) + \hat{\beta}_2(Y_2) + \cdots + \hat{\beta}_m(Y_m)))\right)^{-1}$$

$$= \left(1 + e^{-Y'\underline{\beta}}\right)^{-1}.$$

For a given year and strata (or sub-strata), an estimate of g_0 and its estimated sampling variance can be obtained. Details on calculation of the sampling variance for this approach are known.

5. Sample Size, n and the Spatial Variation in n

Typically, the largest variance component in these surveys is the spatial variation in the number of tortoises detected, var(n). If all tortoises are distributed at random (i.e., Poisson distributed), then $\hat{\text{var}}(n) = n$. From a cursory analysis of some existing line transect data on tortoise, this is extremely unlikely and, instead, we expect var(n) $> n$. Thus the spatial variance must be estimated empirically from the independent lines (l_i), where the total line length $= L = \sum_{i=1}^{} l_i$, then

$$\hat{\text{var}}(n) = L\sum_{i=1}^{k} l_i\left(\frac{n_i}{l_i} - \frac{n}{L}\right)^2 / (k-1),$$

where k is the number of lines (squares). If each line is of equal length, a simpler equation is available, but the one above is appropriate in either case. After several years of data have been collected, some more parsimonious methods of estimating this variance component can be considered (see Buckland et al. 1993:362-364 for an example). Program DISTANCE allows for these more advanced methods.

Eventually, some more detailed sub-stratification might be possible using specific habitat and soil types, elevation, and other variables. This would likely further reduce the spatial variation in numbers detected. In fact, there are several advanced analysis

methods that could be explored in the analysis of these data, once 5-7 years of data are available (colleagues at Oregon State University and the University of St. Andrews in Scotland are currently working on these issues).

6. Estimation of Tortoise Density and Abundance (from the n, $\hat{P}_a$, and $\hat{g}_0$)

Tortoise density (D) each year for each of the strata would be estimated from standard line transect theory,

$$\hat{D} = \frac{n}{2wL \cdot \hat{P}_a \cdot \hat{g}_0},$$

where n is the number of tortoise detected by Survey Team A, P_a is the (average) proportion of the tortoises detected within a transect of width w, and g_0 is the (average) proportion of tortoises that were detected "above ground" during the survey period as estimated by Survey Team B. Of course, $\hat{D}$ is an estimate of the *average* density (both above and below ground), during the time of the survey.

For any particular year and stratum level, the estimated sampling variance of $\hat{D}$ is,

$$\hat{v}ar(\hat{D}) = \hat{D}^2 \cdot \left\{ \frac{\hat{v}ar(n)}{n^2} + \frac{\hat{v}ar(\hat{P}_a)}{(\hat{P}_a)^2} + \frac{\hat{v}ar(\hat{g}_0)}{(\hat{g}_0)^2} \right\}.$$

This expression allows the variance components to be expressed as a percentage and this has implications for the (slight) redesign of the survey after 2-3 years of data are available. Generally, this information would suggest increased effort for either Team A or B (probably at the expense of the other) to improve the overall survey precision. As years progressed, it seems likely that the effort of Survey Team A would be increased relative to the effort of Survey Team B (a trade-off). Program DISTANCE routinely computes all of these quantities, given $\hat{g}_0$ and its estimated standard error.

7. Stratification

Stratification is important to allow estimates of density for each Recovery Unit and each DWMA and to monitor trends in density in each of these areas, as well as the total area samples (the sum of the 6 Recovery Units). Stratification also allows increased efficiency (higher precision) in the estimates of density. Mean density $\hat{D}$ for the total area surveyed is the average of the individual estimates, weighted by the respective stratum areas A_j:

$$\hat{D} = \frac{\sum_j A_j \hat{D}_j}{A}, \text{ with } A = \sum_j A_j.$$

The precision of this estimate is computed by DISTANCE from known formulae. Here, these is considerable flexibility in that some parameters can be pooled across strata in the analysis. Population abundance (N) is estimated by

$$\hat{N} = A \cdot \hat{D} = \sum A_j \hat{D}_i,$$

with

$$\hat{\text{var}}(\hat{N}) = A^2 \cdot \hat{\text{var}}(\hat{D}).$$

For each survey year, estimates of both density (D) and abundance (N) would be available for each primary and secondary strata and for the total area surveyed. Associated estimates of precision would be available (standard errors, coefficients of variation, confidence intervals). If good precision is obtained at the strata level, then the ability to detect meaningful trends might be fairly good. We recognize that special conditions in a particular Recovery Unit might warrant additional sampling effort (and related expenses).

Some additional stratification might be possible after 3-5 years of data are available. This has implications for both Survey Teams A and B in that habitat records must also be recorded at each detection or location (GPS site). Such post-stratification would likely require a GIS system to delineate the size and extent of the new strata types.

8. Annual Rate of Population Change, λ

The finite rate of population change (λ) can be estimated from the estimates of annual population density, either on a stratum basis or for the entire population surveyed. Given $\hat{D}_1, \hat{D}_2, ..., \hat{D}_r$, the estimate of the finite rate of population change for year i is merely,

$$\hat{\lambda}_i = \hat{D}_{i+1}/\hat{D}_i,$$

and, in the simplest case where the annual estimates of density are independent, with

$$\hat{\text{se}}(\hat{\lambda}) = \hat{\lambda}^2 \left((\text{cv}(\hat{D}_{i+1}))^2 + (\text{cv}(\hat{D}_i))^2 \right),$$

where cv is the coefficient of variation. Confidence intervals and other measures of precision on λ could easily be computed. For areas as large as a Recovery Unit, for example, immigration -- emigration would likely be unimportant (e.g., small), thus $\hat{\lambda}$ would validly reflect the finite change (i.e., births -- deaths) as a rate.

After several years of estimates of density were available, one could regress $\log_e(\hat{D})$ on years $(t = 1, 2, 3, ...)$ as

$$\log_e(\hat{D}) = \hat{\alpha} + \hat{r}(t)$$

and compute $\hat{\lambda} = e^{\hat{r}}$ for an estimate of the average value for λ. More advanced approaches are available, but would take us too far astray here. Under these approaches, the sampling variance can be obtained using the delta method.

Several other demographic parameters can be estimated as a result of the monitoring program proposed. For example, unbiased estimates of the population sex ratio can be made as

$$\text{Sex ratio} = \hat{D}_m/\hat{D}_f,$$

where estimated mean density of tortoise of a certain size class of interest would be used

in the computations. [Note, the estimator n_m/n_f would be a biased estimator of the sex ratio and we would not recommend it.] Improved estimates of density of tortoise by size class can be made by using tortoise size as a covariate in DISTANCE. Size-specific survival probabilities could also be estimated and estimates of individual growth rate could be made from the sample of radio-marked tortoises as an indirect result of the work by Survey Team B. These parameters could be best estimated from the total of the 6 Recovery Units, however other partitions of the data could be used, particularly after several years of monitoring data are available.

9. Other Considerations and Advanced Methods

After several years of good survey data are available then some advanced analysis options become available. For example, possible pooling of the distance data to obtain a pooled estimate of the probability of detection, P_a; this would increase precision, but make the annual estimates to be dependent. Gilbert et al. (1996) provides an example of a long-term monitoring program using line transects sampling where such pooling was effective. A model of g_0 could be used, rather than annual estimation of this parameter from the data collected by Survey Team B. Options for the analysis of these alternatives are in program DISTANCE (including the bootstrap and various pooling strategies). State-of-the-art methods for detection of time trends will almost certainly be available in DISTANCE long before the long-term data from this survey are available.

Program DISTANCE routinely computes parameter estimates, goodness of fit statistics, graphs of the distance data vs. the estimated detection function, estimates of precision, model fitting alternatives, and other useful information. The program is available without charge and runs well on a 486 or Pentium machine under either DOS or *Windows*. There may be a *Windows* interface for DISTANCE within a year. This would allow the investigator to respond to questions asked by the software – this would make the program easy to use for the beginner. Investigators, researchers, managers could obtain the software and explore the data as they wished.

Program DISTANCE would need to be modified in two ways: (1) to allow individual estimates of g_0 and its standard error at the stratum level, and (2) to allow sub-strata within a strata (e.g., several Desert Wildlife Management Areas within a Recovery Unit). These modifications would not need to be implemented for 2-3 years after the survey was initiated.

After about 5-7 years the recapture data on individually marked tortoise could be summarized as a capture history matrix and analyzed using an open population capture-recapture model. This would allow estimates of annual apparent survival probabilities by sex and size (Burnham et al. 1987, Lebreton et al. 1992, Anderson et al. 1995). If recapture probabilities were reasonably high, these estimates would also be quite useful in examining trends in survival and assessing delisting criteria. An example might be that the transect surveys indicated a decreasing population; then if the estimated survival probabilities showed no change, one might conclude that poor recruitment was responsible for the decline. In addition, estimates of annual survival probabilities by sex and size could be made from the telemetered tortoises if radios of a special type were to be deployed.

Dead tortoises found on transects would be recorded, perpendicular distance from the transect measured and removed from the area searched. A multi-year analysis of these data would yield some independent insights into possible trends in mortality.

Some consideration should be given to adaptive sampling. Such designs are effective for populations that are more spatially aggregated than desert tortoise, however, they might still be considered. The "square" transects suggested here would have to be altered in some way. Generally, we see only slight advantages in the adaptive designs for the desert tortoise and will not be considered further here.

10. The Power to Detect Trends

The monitoring program would be designed to detect trends in tortoise density and abundance. One objective is to detect a drastic, short-term (4-5 years) decline, while a second objective is to detect small, positive trends over long time periods (20-25 years). Theory to assess the "power" to detect such trends has been implemented in program TRENDS (see, e.g., Gerrodette 1987, 1991, Link and Hatfield 1990). Statistical power to detect trends of hypothetical magnitudes is a somewhat useful concept for planing purposes, but hinges on the selection of an arbitrary α level (i.e., power can always be increased by simply picking a larger value of α: e.g., 0.20 or 0.15 instead of the usual 0.05 or 0.01). Further, the natural, annual variability in density (in addition to any trend) is important in computing power and this quantity is difficult to specify in advance. We explored the subject of power using program TRENDS to gain some rough insights into what might be expected from the overall design we are recommending.

Drastic, Short-term Decline. – Assuming a cv = 0.15 and α = 0.15, the survey could detect a 12% annual decline in 4 years of survey data with power = 0.78 and the power increased to 0.97 with an additional year of survey data.

Long-term Increases. – Assuming a cv = 0.15 and α = 0.15, the survey could detect a 2% annual increase in 25 years of survey data with power = 1.0 and a 1% annual increase with power = 0.86.

In both cases, the power to detect trends for an individual Recovery Unit would be lower. For example, if cv = 0.35 and α = 0.15, the power to detect a 2% annual increase over 25 years would drop to 0.72.

Estimates of power drop off slowly if sampling is done only every second or third year. For example, if the population is increasing at 2% per year, the power is 1.0 for annual sampling, 0.99 for biennial sampling, 0.92 for triennial sampling, and then drops to 0.83 for sampling at only 4 year intervals (again, assuming cv=0.15 and α=0.15). These estimate support recommendations concerning biennial or triennial sampling programs, once a firm baseline has been established. This subject would have to receive further scrutiny if a detailed design was attempted.

11. Sample Sizes and Costs

Large sample sizes will be required and costs can be expected to be significant (and can be estimated with decent accuracy). Chapter 7 of Buckland et al. (1993) provides formulae to estimate sample sizes, costs, precision, etc. Adequate data exist (e.g., encounter rates n_i/L_i, for several strata i) to allow a careful survey design to be developed. This would be the next step, if there is a serious effort to begin a monitoring program for the desert tortoise.

From the past data available, the coefficient of variation of estimated density can be computed empirically and denoted as $cv(\hat{D})$. Then line length required (say L),

and thus sample size, to achieve the target (the subscript t is used to denote this target level of precision) precision is given by

$$L = \frac{L_0 (cv(\hat{D}))^2}{(cv_t(\hat{D}))^2} \, ,$$

where L_0 is the actual length of line used to obtain the estimate, $\hat{D}$. This would be a useful starting point at the stratum level, although some level of approximation would be required for the variance component for g_0. For example, for planning purposes,

$$\mathrm{var}(g_0) = 2g_0(1-g_0)/n_T$$

might be useful for, at least, the first (pilot) year of survey.

For fixed total line length L_T, it is possible to obtain near optimal, total line length for each of the strata (i), expressible as the ratios

$$\frac{L_i}{L} = \frac{A_i \sqrt{D_i}}{\sum A_i \sqrt{D_i}} \, .$$

Here several years of data would provide estimates of average density and then some reallocation of sampling effort could be done for future years of survey. If there was particular interest in one Recovery Unit, then additional sampling effort could be added, but not at the expense of the other Units.

Enough data exist to justify the *a priori* use of a fixed transect width (w) – approximately 30m. Distances would ideally be recorded "exactly" (thus, "ungrouped" data). There should be at least 35 lines (squares) in each Recovery Unit each year. Some desired level of precision should be provided by managers on a strata basis; this is needed in determining total line length by strata (L_i). If individual estimates with high precision are required on individual DWMAs, then sample size requirements (and associated line length required) will increase sharply. There are trade-offs between what might be wanted and what might be fiscally possible (particularly at the DWMA level).

Total line length and expected sample size required remain to be computed. The desired level of precision at the strata level remain to be specified. At this point, an encounter rate (i.e., n/L) of about 0.3/km seems reasonable (based very crudely on some data provided by Drs. Corn and Freilich). This would give an expected number of tortoise detections of about 1.2 per transect (square); however, this would likely be highly variable.

It is interesting to note that 1-ha removal plots have an encounter rate (e.g., number of tortoise found per person day) that is about one half that of line transect sampling. Therefore, this alternative is inefficient and, thus, "costly." The reason for this is interesting. In an optimal removal survey all the objects are found on the first occasion, the second sampling occasion merely provides evidence that all objects were found on the first occasion. However, in a broader context, effort expended on the second occasion produces no data! In addition, the issue of tortoises remaining below ground during the survey is also a problem with the 1-ha removal plots (but could be resolved using a Survey Team B, as proposed here).

We can offer a general order of magnitude for the extent of the survey effort for the reserved areas. These are based on some analysis of past data and some planning done using program DISTANCE. At this early point, we are thinking of using about 300 line transects each year, each about 4 km in length and expecting to detect about 300-400 tortoises in an average year. Assuming >0.3 tortoise detected per km of transect, the coefficient of variation (cv) for the estimated average density for the total reserved area would be approximately 10-15%. The cv for individual primary strata might range from 20 to 40%. These are very preliminary, and include only the precision related to tortoises above ground, but are offered to give readers a rough idea of survey effort needed. Clearly, this is a subject needing more definitive work.

Very preliminary costs estimates, presented at the November, 1996 workshop in Nevada, suggest a figure in the $750-850,000 range for the first year of monitoring. This does not include salaries of permanent people. Detailed computations must be done before a final estimate of first-year costs can be made. Survey costs will tend to decrease in years 2-5, and then decrease again in years > 5. Costs will be high the first year and diminish while, in sharp contrast, the value of the data (the information) accumulates through time.

12. Some Practicalities

The first year should be considered a full-scale pilot survey. The survey results should receive outside, critical, independent review as this would allow some fine-tuning of the survey. The design should be published and a panel should be appointed to enforce standards, review results, and make recommendations. Quality control could be assessed and further training and enhancements implemented. This would be a time to reaffirm management commitment to a long-term survey/monitoring program as improved cost and labor estimates would be available, as well as estimates of precision. Inconsistent funding has hampered previous attempts at monitoring this species and every effort should be made to assure adequate funding over a long time horizon.

We recommend that a full-scale, annual survey be conducted for the first 5 years, followed by biannual surveys of half of the secondary strata in odd-numbered years and biannual surveys of the other half of the secondary strata in even-numbered years. This design would reduce costs considerably, while still having a firm 5 year *baseline* for future trend assessment. Annual monitoring of all strata levels for the first 5 years will provide high power to detect any (feared) drastic, short-term declines in density. Then, the reduced intensity of the sampling program after the 5th year will provide cost-effective monitoring of (expected) increasing trends on the reserved areas over longer time frames. This 2-phased sampling schedule would allow competent field crews to work each year and provide adequate data for both short- and long-term monitoring. There are several important details that will warrant careful scrutiny and refinement after the first 2-4 years of survey data have been collected and carefully analyzed.

The survey/monitoring program for non-reserved lands would be similar to that outlined above. Here, we suggest definition of many large blocks of non-reserved lands somewhat "adjacent" to the reserved lands. A random sample of these adjacent blocks would allow a "pairing" of, say, 10 reserved and non-reserved areas. Given such data, a wide array of effective analysis options exist to compare trends in tortoise densities between reserved and non-reserved areas. An *experiment* to compare density on and off reserved areas is not possible, instead, only a rough comparison of trends can be considered. This entire subject needs much more thought; however, the needed resources will be of the same order of magnitude as that for the reserved areas (described herein).

13. Training – Field Measurements

The field protocol is very important in conducting line transect surveys. The focus must be directed toward assuring the validity of the 2 key assumptions underlying the method. In particular, accurate distances must be obtained – this makes placement and clear marking of the centerline critical. Search behavior must assure that all tortoises on or near the centerline are detected with certainty (if they are on the surface). Several things should be done to increase the integrity and quality of the field surveys. For example, Survey Team A could be told that "model" tortoises had been placed on some of the transect lines (e.g., under dense brush), thus they would be "checked" for the quality of their work (similar to that in the 1-ha removal plots where one team searches for tortoise on the first occasion and their work is "checked" by an independent team during the second occasion). Other incentives could be built into the line transect surveys and similar "checks" could be included in the radio telemetry work done by Survey Team B. Intensive training of teams of observers will be essential (of course, the people in Teams A and B would require very different training). It might be advantageous to rotate some crew members between A and B teams to allow experience and training in the overall survey method.

Unlike plot-type surveys, data from line transect surveys can reveal poor field methods. Thus, problems in the distance data can point to field crews that are sloppy or inadequately trained. For example, if the transect centerline is poorly marked, many detections will often be recorded at 0 distance (called "heaping at zero"). If measurements to detected tortoises are not made accurately, there is a tendency to heap at "round" numbers, such as 5, 10 20, etc; again, these anomalies can be seen in the distance data. Often, there is heaping at w – here tortoises detected just outside the boundary (w) are erroneously included as if they were at distance w. These inaccuracies are revealed by examination of the histrograms of the distance data.

A major workshop and field exercise could be considered before an integrated survey is initiated. This would focus on an understanding of the methodology and the field protocol required. In particular, search behavior and the key assumptions would need considerable emphasis. The monitoring methodologies suggested here will almost certainly fail if untrained observers conduct the survey. Novice observers and volunteers might have a place in Survey Team B, but only if working under the careful direction of a trained observer. In addition, such novice observers could serve as an extra on Survey Team B. Good line transect protocol in the field does not just "happen," adequate care, close supervision, and some "checks" must be done during data collection.

If crew members have a good understanding of distance sampling, they can help each other perform well in the field. For example, it is easier to detect tortoises on transects when the observer is walking uphill, as opposed to downhill. This fact should bear on how the legs of each line (square) are to be covered during the survey. Each crew is to cover only one relatively short (e.g., 4.0 km) line per day; thus, there is time to do very high quality data collection.

14. Record Keeping, Data Entry and Repository

Ideally, observers would enter data into a computerized medium at the end of each day, checking everything carefully against the field forms. Further rechecking and backup should occur at the end of each week. A protocol would need to be developed to handle the data during the field season. One person should be assigned as a curator of these important records and the original field form and electronic files should be kept in

an appropriate repository. Long-term monitoring must place a high premium on past data, its accuracy and safe storage. Thorough documentation, including maps showing stratum boundaries, location of sampling units, must be kept and available. Responsibility must be clearly assigned for these important tasks.

15. Perspectives

We believe the survey should focus on unbiased and precise estimates of population density (D) and abundance (N) and trends in these parameters over time, by primary strata and for the total area surveyed. It is time to put aside notions of an annual "index" to population density (or size) and trends in such an "index." It seems fruitless to record scat, burrows of various types, or scratch marks on rocks in the blind hope that these are consistently, linearly related to the parameter of interest. Further, it seems impossible that as both the environment and the size of the tortoise population change considerably over long time frames, the index retains its original (but unknown) relationship to the parameters of interest. The misplaced notion of an "index" has not worked for other species and does not permit valid inference concerning the status of populations. Similarly, use of merely the number of tortoises somehow detected as an "index" to density is similarly without validity and should not receive consideration.

We must note that the word *census* refers to a *total enumeration* of a population (see any dictionary; the misuse of the word "census" seems to arise from the amateur birders). Thus, during a census, one counts the members of the population "one by one" until the final member is counted (i.e., 1, 2, 3, ..., N). The population size is then just N; it has no sampling variance, because no *sampling* was done! If a census of desert tortoise is feasible, it would certainly be the preferred approach. In reality, a *census* is simply not possible, thus a *survey* should be considered, whereby a (probabilistic) *sample* of areas is taken and inductive inferences made about the population parameter (N or D), based on the information in the *sample.*

16. Acknowledgments

17. Literature Cited

Anderson, D. R., G. C. White, and K. P. Burnham. 1995. Some specialized risk assessment methodologies for vertebrate populations. Environmental and Ecological Statistics 2:91-115.

Buckland, S. T., D. R. Anderson, K. P. Burnham, and J. L. Laake. 1993. Distance sampling: estimating abundance of biological populations. Chapman and Hall, London. 446pp.

Burnham, K. P., D. R. Anderson, G. C. White, C. Brownie, and K. H. Pollock. 1987. Design and analysis methods for fish survival experiments based on release-recapture. American Fisheries Society Monograph, 5, 437pp.

Gerrodette, T. 1987. A power analysis for detecting trends. Ecology 68:1364-1372.

Gerrodette, T. 1991. Models of power of detecting trends – a reply to Link and Hatfield. Ecology 72:1889-1892.

Gilbert, D. W., D. R. Anderson, J. R. Ringelman, and M. R. Szymczak. 1996. Response of nesting ducks to habitat and management on the Monte Vista National Wildlife Refuge, Colorado. Wildlife Monographs, 131:1-44.

Laake, J. L., S. T. Buckland, D. R. Anderson, and K. P. Burnham. 1994. DISTANCE User's Guide. Version 2.1. Colorado Cooperative Fish and Wildlife Research Unit, Colorado State University, Fort Collins, CO 84pp.

Lebreton, J-D., K. P. Burnham, J Clobert, and D. R. Anderson. 1992. Modeling survival and testing biological hypotheses using marked animals: A unified approach with case studies. Ecological Monographs, 62:67-118.

Link, W. A., an J. S. Hatfield. 1990. Power calculations and model selection for trend analysis: a comment. Ecology 71:1217-1220.

December 10, 1996

Example Field Form

Strata Name _   Year _ _ _ _ _ _ _ _

Temperature at noon (C°) _ _ _ _ _   Cloud cover (%) _ _ _ _ _   Elevation (m) _ _ _ _ _ _

Observer names _ _ _ _ _ _ _ _ _   _ _ _ _ _ _ _ _ _ _   Date _ _ _ _ _ _ _ _ _ _ _ _

Line number _ _ _ _   Start time _ _ _ _   End time _ _ _ _   Starting GPS location _ _ _ _

Detection number	Perpendicular distance (m)	Length (cm)	Sex	GPS location	Tag number	Habitat type	Soil type
1	_ _ _ _	_ _	_ _	_ _	_ _	_ _	_ _
2	_ _ _ _	_ _	_ _	_ _	_ _	_ _	_ _
3	_ _ _ _	_ _	_ _	_ _	_ _		
⋮							

[Note, if no detections were made on line i, then the count for that line is 0]

A field form would be completed for each survey line, l_i. Notes should be taken regarding any tag number noted or the number of the new tag attached upon detection. Grazing history could be noted for each transect. Details of the field form need development, this example gives only the most minimal information. Strict handling protocol must be included in training and during survey conduct.